JOURNAL OF COMPLEXITY 5, 466–488 (1989)

The Complexity of Dynamic Programming*,†

CHEE-SENG CHOW AND JOHN N. TSITSIKLIS

Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received April 3, 1989

We provide tight lower bounds on the computational complexity of discrete-time, stationary, infinite horizon, discounted stochastic control problems, for the case where the state space is continuous and the problem is to be solved approximately, within a specified accuracy. We study the dependence of the complexity on the desired accuracy and on the discount factor. © 1989 Academic Press, Inc.

1. INTRODUCTION

This paper addresses issues related to the computational complexity of solving discrete-time stochastic control problems defined on a continuous state space. There has been a great deal of research on the computational aspects of stochastic control when the state space is finite (Bertsekas, 1987; Papadimitriou and Tsitsiklis, 1987; Hartley *et al.*, 1980). However, much less has been accomplished for the case of continuous state spaces. An explanation for this state of affairs could be that such problems are very demanding computationally, with most realistic problems lying beyond the capabilities of commercial computers. However, with advances in computer hardware and with the availability of new powerful architectures, it is to be expected that the numerical solution of continuous-state stochastic control problems will become much more common, hence the motivation for our work.

Let S be some subset of $\Re^n$ which is the state space of a controlled stochastic process. A large class of discrete-time stochastic control prob-

* Invited Paper.

† Research supported by the NSF under Grant ECS-8552419, with matching funds from Bellcore and Du Pont, and by the ARO under Grant DAAL03-86-K-0171.

0885-064X/89 $3.00

lems boils down to the computation of a fixed point J^* of the nonlinear operator T (acting on a space of functions on the set S) defined by

$$(TJ)(x) = \inf_{u \in U} \left[g(x, u) + \alpha \int_S J(y) P(y|x, u)\, dy \right], \qquad \forall x \in S. \quad (1.1)$$

Here, $U \subset \Re^m$ is the control space, $g(x, u)$ is the cost incurred if the current state is x and control u is applied, $\alpha \in (0, 1)$ is a discount factor, and $P(y|x, u)$ is a stochastic kernel that specifies the probability distribution of the next state y, when the current state is x and control u is applied. Then, $J^*(x)$ is interpreted as the value of the expected discounted cost, starting from state x, and provided that the control actions are chosen optimally (see Section 2).

A fixed point J^* of the operator T cannot be determined analytically except for a limited class of examples. On the other hand, an approximation to such a fixed point can be computed by suitably discretizing the state and control spaces, and then solving a finite-dimensional version of the problem. There has been some work on such discretization methods, with typical results demonstrating that as the discretization becomes finer, the resulting approximation of J^* becomes more and more accurate (Whitt, 1978; Bertsekas, 1975; Chow and Tsitsiklis, 1989). Upper bounds on the approximation error are also available. One of the consequences of the results to be derived in the present paper is that some of the earlier upper bounds are tight within a constant factor.

Once the original problem is discretized, there is a choice of numerical methods for solving the discrete problem (Bertsekas, 1987). One particular choice is studied in a companion paper (Chow and Tsitsiklis, 1989), where it is shown that the total computational effort is closely related to the amount of work needed in discretizing the problem, when a suitable multigrid method is employed. Thus, the results in the present paper demonstrate that the algorithm of Chow and Tsitsiklis (1989) is close to optimal (and sometimes optimal) as far as its complexity is concerned.

In the special case where the control space U consists of a single element, the minimization in Eq. (1.1) is redundant, and the fixed point equation $J^* = TJ^*$ becomes a (linear) Fredholm equation of the second kind. Thus, the results in the present paper, in conjunction with the algorithms of Chow and Tsitsiklis (1989), characterize the computational complexity of the approximate solution of Fredholm equations of the second kind. There has been some prior research on this subject (Werschulz, 1985). Our work is different in a number of respects that are discussed in Section 4.

The paper is organized as follows. In Section 2, we introduce our assumptions, define the problem to be solved, and state its relation to stochastic control. We also outline the model of computation to be employed. Finally, we state some known upper bounds on the approximation error introduced by the discretization of the problem, and translate them to complexity bounds. In Section 3, which contains our main results, lower bounds are derived on the complexity of our problem. Finally, Section 4 contains our conclusions and some discussion of related issues.

2. Problem Definition

Let $S = [0, 1]^n$ and $U = [0, 1]^m$. Let α be a scalar belonging to $(0, 1)$. Let $g: S \times U \mapsto \Re$ and $P: S \times S \times U \mapsto \Re$ be some functions. Let K be some positive constant.

Assumption 2.1: (a) $0 \leq P(y|x, u)$, $\forall y, x \in S$, $\forall u \in U$.

(b) $|P(y|x, u) - P(y'|x', u')| \leq K\|(y, x, u) - (y', x', u')\|_\infty$, $\forall y, y', x, x' \in S$, $\forall u, u' \in U$.

(c) $|g(x, u) - g(x', u')| \leq K\|(x, u) - (x', u')\|_\infty$, $\forall x, x' \in S$, $\forall u, u' \in U$.

(d) $|g(x, u)| \leq K$, $\forall x \in S$, $\forall u \in U$.

Assumption 2.2. $\int_S P(y|x, u)\, dy = 1$, $\forall x \in S$, $\forall u \in U$.

According to Assumptions 2.1 and 2.2, for any fixed $x \in S$ and $u \in U$, the function $P(\cdot|x, u)$ is a probability density on the set S. Furthermore, P and g are Lipschitz continuous with Lipschitz constant K.

Let $\mathscr{C}(S)$ be the set of all continuous real-valued functions on the set S. We define the operator $T: \mathscr{C}(S) \mapsto \mathscr{C}(S)$ by letting

$$(TJ)(x) = \min_{u \in U} \left[g(x, u) + \alpha \int_S J(y) P(y|x, u)\, dy \right], \qquad \forall J \in \mathscr{C}(S),\ \forall x \in S. \tag{2.1}$$

(The fact that T maps $\mathscr{C}(S)$ into itself is proved by Chow and Tsitsiklis, 1989). The space $\mathscr{C}(S)$ endowed with the norm $\|J\|_\infty = \max_{x \in S}|J(x)|$ is a Banach space. Furthermore, T has the contraction property

$$\|TJ - TJ'\|_\infty \leq \alpha \|J - J'\|_\infty, \qquad \forall J, J' \in \mathscr{C}(S). \tag{2.2}$$

Since $\alpha \in (0, 1)$, T is a contraction operator, and therefore has a unique fixed point $J^* \in \mathscr{C}(S)$. The equation $TJ = J$, of which J^* is the unique solution, is known as *Bellman's equation* or as the *dynamic programming* *equation*. We are interested in the computational aspects of the approximate evaluation of J^*.

Stochastic Control Interpretation

Let Π be the set of all Borel measurable functions $\mu: S \mapsto U$. Let Π^∞ be the set of all sequences of elements of Π. An element $\pi = (\mu_0, \mu_1, \ldots)$ of Π^∞, called a *policy* (also known as *feedback law* or *control law*), is to be viewed as a prescription for choosing an action $\mu_t(x_t) \in U$ at time t, as a function of the current state x_t of a controlled stochastic process. More precisely, given a policy $\pi = (\mu_0, \mu_1, \ldots)$, we define a (generally, nonstationary) Markov process $\{x_t^\pi \mid t = 0, 1, \ldots\}$ on the state space S by letting $P(\cdot|x_t^\pi, \mu_t(x_t^\pi))$ be the probability density function of x_{t+1}^π, conditioned on x_t^π. The cost $J_\pi(x)$ associated to such a policy is defined (as a function of the initial state x) by

$$J_\pi(x) = E\left[\sum_{t=0}^{\infty} \alpha^t g(x_t^\pi, \mu_t(x_t^\pi)) \mid x_0^\pi = x\right]. \tag{2.3}$$

[Note that the infinite sum is absolutely convergent and bounded by $K/(1-\alpha)$ because $\alpha \in (0, 1)$ and the function g is bounded by K.] For any $x \in S$, we define $\hat{J}(x)$ by letting

$$\hat{J}(x) = \inf_{\pi \in \Pi^\infty} J_\pi(x), \tag{2.4}$$

and this defines a function $\hat{J}: S \mapsto \Re$. This function, known as the *cost-to-go* or *value* function, represents the least possible cost as a function of the initial state of the controlled process. A policy $\pi \in \Pi^\infty$ is called *optimal* if $J_\pi(x) = \hat{J}(x)$ for all $x \in S$. The central result of dynamic programming[1] states that $\hat{J}$ coincides with the fixed point J^* of the operator T. Furthermore, once J^* is available, it is straightforward to determine an optimal policy. [This is done as follows: Consider Eq. (2.1) with J replaced by J^*. For each $x \in S$, choose some u that attains the minimum in Eq. (2.1), and let $\mu_t(x) = u$ for each t.] This justifies our interest in the function J^*.

The case where the discount factor α approaches 1 from below is of substantial theoretical and practical interest. For example, as $\alpha \uparrow 1$, one obtains, in the limit, the solution to a certain "average cost problem" (Bertsekas, 1987). Also, if one deals with a discounted continuous-time stochastic control problem and the time step is discretized, one obtains a discrete-time discounted problem in which the discount factor approaches 1 as the time discretization step is made smaller.

[1]This result requires certain technical assumptions. Assumption 2.1 turns out to be sufficient (Bertsekas and Shreve, 1978; Chow and Tsitsiklis, 1989).

In practical stochastic control problems, the state and control spaces could be arbitrary subsets of $\Re^n$ and $\Re^m$, respectively, and there could be state-dependent constraints on the allowed values of u. Such problems can only be harder than the special case studied here. Thus, the lower bounds to be derived in Section 3 apply more generally. A similar comment applies to the smoothness conditions on g and P that have been imposed in Assumption 2.1.

Further Assumptions

In the case of stochastic control problems defined on a finite state space it is known that the convergence of certain algorithms for computing J^* is much faster when the controlled process satisfies certain *mixing conditions* (Bertsekas, 1987) and the same is true in our case as well (Chow and Tsitsiklis, 1989). Our next assumption introduces a condition of this type.

Assumption 2.3. There exists a constant $\rho > 0$ such that

$$\int_S \min_{x \in S,\, u \in U} P(y|x, u)\, dy \geq \rho. \tag{2.5}$$

Intuitively, Assumption 2.3 states that no matter what the current state is and what control is applied, there are certain states in S (of positive Lebesgue measure) for which the probability density of being visited at the next time step is positive. This ensures that the effects of initial conditions are quickly washed out.

In an alternative class of stochastic control problems, at any given time there is a certain probability, depending on the current state and the control being applied, that the process is terminated and costs stop accruing. Such a formulation is captured by allowing $P(\cdot|x, u)$ to be a subprobability measure, as in the following assumption:

Assumption 2.4. $\int_S P(y|x, u)\, dy \leq 1$, $\forall x \in S$, $\forall u \in U$.

Model of Computation

Our computational task is completely determined by the functions P and g, the discount factor α, and the desired accuracy ε. Accordingly, a tuple $(P, g, \alpha, \varepsilon)$ will be called an *instance*. We then define a *problem* as a class of instances. In our context, different problems will correspond to different choices of assumptions.

In order to talk meaningfully about the approximate computation of J^*, we need a suitable model of computation. We use a real-number model of computation (Traub *et al.*, 1988; Nemirovsky and Yudin, 1983) in which a processor:

(a) Performs comparisons of real numbers or infinite precision arithmetic operations in unit time.

(b) Submits queries $(y, x, u) \in S \times S \times U$ to an "oracle" and receives as answers the values of $g(x, u)$ and $P(y|x, u)$. [We then say that the processor *samples* (y, x, u).] Queries can be submitted at any time in the course of the computation and this allows the values of (y, x, u) in a query to depend on earlier computations or on the answers to earlier queries. (We will therefore be dealing with "adaptive" algorithms, in the sense of Traub *et al.*, 1988.)

An algorithm in the above model of computation can be loosely defined as a program that determines the computations to be performed and the queries to be submitted. An algorithm is said to be *correct* (for a given problem) if for every instance $(P, g, \alpha, \varepsilon)$ of the problem, it outputs a piecewise constant function J such that $\|J^* - J\|_\infty \leq \varepsilon$, in some prespecified format. A natural format for the representation of the output is as follows. The processor outputs a parameter h that signifies that the state space S has been partitioned into cubes of volume h^n, and then outputs the value of J on each one of these cubes, with the understanding that J is constant on each one of these cubes.

The complexity of an algorithm of above described type is defined as the sum of:

(a) the number of oracle queries;

(b) the number of arithmetic operations performed by the algorithm.

A very fruitful method for establishing lower bounds on the complexity of any algorithm consists of lower bounding the number of queries that must be submitted for the desired accuracy to be attainable. The typical argument here is that if the number of queries is small then the available information on the problem being solved is insufficient.

Let the dimensions n, m of the state and control spaces be fixed and let us view the constants K and ρ of Assumptions 2.1–2.3 as absolute constants. We consider three different problems:

(a) Problem $\mathcal{P}_{\text{prob}}$, which consists of all instances that satisfy Assumptions 2.1 and 2.2. [In particular, $P(\cdot|x, u)$ is a probability measure for all (x, u).]

(b) Problem $\mathcal{P}_{\text{mix}}$, which consists of all instances that satisfy Assumptions 2.1–2.3. (That is, a mixing condition is also in effect.)

(c) Problem $\mathcal{P}_{\text{sub}}$ which consists of all instances that satisfy Assumptions 2.1, 2.3–2.4. [That is, the mixing condition is still in effect, but $P(\cdot|x, u)$ is only a subprobability measure.]

Let us fix some ε and α. We define $\mathbf{C}_{\text{prob}}(\alpha, \varepsilon)$ as the minimum (over all correct algorithms for the problem $\mathcal{P}_{\text{prob}}$) of the number of queries, in the worst case over all instances $(P, g, \alpha, \varepsilon)$ of $\mathcal{P}_{\text{prob}}$. The quantities $\mathbf{C}_{\text{mix}}(\alpha, \varepsilon)$ and $\mathbf{C}_{\text{sub}}(\alpha, \varepsilon)$ are defined similarly, by replacing problem $\mathcal{P}_{\text{prob}}$ by $\mathcal{P}_{\text{mix}}$ and $\mathcal{P}_{\text{sub}}$, respectively.

The following upper bounds, together with discretization procedures that stay within these bounds, can be found in Whitt (1978) and Chow and Tsitsiklis (1989).

$$\mathbf{C}_{\text{prob}}(\alpha, \varepsilon) = O\left(\frac{1}{((1-\alpha)^2\varepsilon)^{2n+m}}\right), \tag{2.6}$$

$$\mathbf{C}_{\text{mix}}(\alpha, \varepsilon) = O\left(\frac{1}{((1-\alpha)\varepsilon)^{2n+m}}\right), \tag{2.7}$$

$$\mathbf{C}_{\text{sub}}(\alpha, \varepsilon) = O\left(\frac{1}{((1-\alpha)^2\varepsilon)^{2n+m}}\right). \tag{2.8}$$

We have used the $O(\cdot)$ notation, which should be interpreted as follows. Let f and h be functions from $(0, 1) \times (0, \infty)$ into $[0, \infty)$. We write $f(\alpha, \varepsilon) = O(h(\alpha, \varepsilon))$ if there exist constants $c > 0$, $\varepsilon_0 > 0$, and $\alpha_0 \in (0, 1)$ such that $f(\alpha, \varepsilon) \leq ch(\alpha, \varepsilon)$ for all $\varepsilon \in (0, \varepsilon_0)$ and all $\alpha \in (\alpha_0, 1)$. (These constants are allowed to depend on the absolute constants n, m, K, *and* ρ.) Later, we will aslo use the notation $f(\alpha, \varepsilon) = \Omega(h(\alpha, \varepsilon))$, which is equivalent to $h(\alpha, \varepsilon) = O(f(\alpha, \varepsilon))$.

3. Lower Bounds

In this section, we prove that the upper bounds of Eqs. (2.6)–(2.8) are tight, by establishing the corresponding lower bounds. Our results rest on an "adversary" argument that is commonly used to establish lower bounds on the number of queries an algorithm must make. The outline of the argument is as follows. Suppose that a certain algorithm makes at most A queries. We consider a particular instance $(P, g, \alpha, \varepsilon)$ and we let X be the set of triples (y, x, u) sampled by the algorithm when presented with that instance. We then construct an alternative instance $(\hat{P}, \hat{g}, \alpha, \varepsilon)$ such that $P(y|x, u) = \hat{P}(y|x, u)$ and $g(x, u) = \hat{g}(x, u)$ for all $(y, x, u) \in X$. The algorithm has no means of distinguishing between the two instances and must produce the same output J in both cases. Let J^* and $\hat{J}^*$ be the optimal cost functions for the two problems. If we can manage so that $\|J^* - \hat{J}^*\|_\infty > 2\varepsilon$, then at least one of the inequalities $\|J - J^*\|_\infty > \varepsilon$ and $\|J - \hat{J}^*\|_\infty > \varepsilon$ must hold. It follows that the algorithm cannot succeed for all problem instances, and therefore the least number of queries necessary for the problem is larger than the cardinality of A.

THEOREM 3.1 (lower bound under Assumptions 2.1–2.3). *For any $K > 0$, $\rho \in (0, 1)$, m, and n, we have*

$$\mathbf{C}_{\text{mix}}(\alpha, \varepsilon) = \Omega\left(\frac{1}{((1-\alpha)\varepsilon)^{2n+m}}\right).$$

Proof. We only prove the result for the case $K = 1$ and $\rho = 1/2$. The proof for the general case is identical except for a minor modification discussed at the end of the proof. Let us also fix the dimensions m, n of the problem. Throughout the proof, an *absolute constant* will stand for a constant that can depend only on m, n, not on any other parameters.

We fix some $\varepsilon > 0$ and some $\alpha \in (1/2, 1)$. Let us consider some algorithm that is correct for the problem $\mathcal{P}_{\text{mix}}$ and suppose that the number of queries is at most A for every instance with those particular values of α and ε. We will derive a lower bound on A.

We choose a positive scalar δ so that $1/\delta$ is an integer multiple of 16 and such that

$$\frac{1}{\delta_0} \leq \frac{1}{\delta} \leq \frac{1}{\delta_0} + 16, \tag{3.1}$$

where δ_0 satisfies

$$A = \frac{1}{4\delta_0^{2n+m}}. \tag{3.2}$$

We partition the set $S \times S \times U$ into cubic cells of volume δ^{2n+m}. (In particular, there will be $1/\delta^{2n+m}$ cells.) This is done by first specifying the "centers" of the cells. Let $\tilde{S}$ be the set of all $x = (x_1, \ldots, x_n) \in S$ such that each component x_i is of the form $x_i = (t + (1/2))\delta$, where t is a nonnegative integer smaller than $1/\delta$. Similarly, we let $\tilde{U}$ be the set of all $u = (u_1, \ldots, u_m) \in U$ such that each u_i is of the form $u_i = (t + (1/2))\delta$, where again t is a nonnegative integer smaller than $1/\delta$. For any $(\tilde{y}, \tilde{x}, \tilde{u}) \in \tilde{S} \times \tilde{S} \times \tilde{U}$, we define the *cell* $C_{\tilde{y},\tilde{x},\tilde{u}}$ by letting

$$C_{\tilde{y},\tilde{x},\tilde{u}} = \left\{(y, x, u) \in S \times S \times U \;\middle|\; \|(y, x, u) - (\tilde{y}, \tilde{x}, \tilde{u})\|_\infty < \frac{\delta}{2}\right\}.$$

Clearly, the cardinality of $\tilde{S}$ and $\tilde{U}$ is $1/\delta^n$ and $1/\delta^m$, respectively. It follows that there is a total of $1/\delta^{2n+m}$ cells. Note that distinct cells are disjoint.

For any $(\tilde{y}, \tilde{x}, \tilde{u}) \in \tilde{S} \times \tilde{S} \times \tilde{U}$, we define a function $E_{\tilde{y},\tilde{x},\tilde{u}}: S \times S \times U \mapsto \mathfrak{R}$, by letting

$$E_{\tilde{y},\tilde{x},\tilde{u}}(y|x, u) = 0, \qquad \text{if } (y, x, u) \notin C_{\tilde{y},\tilde{x},\tilde{u}},$$

$$E_{\tilde{y},\tilde{x},\tilde{u}}(y|x, u) = \frac{\delta}{2} - \|(y, x, u) - (\tilde{y}, \tilde{x}, \tilde{u})\|_\infty, \qquad \text{if } (y, x, u) \in C_{\tilde{y},\tilde{x},\tilde{u}}. \tag{3.3}$$

Thus, $E_{\tilde{y},\tilde{x},\tilde{u}}$ is just a "pyramid" of height $\delta/2$ whose base is the cell $C_{\tilde{y},\tilde{x},\tilde{u}}$. The triangle inequality applied to the norm $\|\cdot\|_\infty$ shows that

$$|E_{\tilde{y},\tilde{x},\tilde{u}}(y|x, u) - E_{\tilde{y},\tilde{x},\tilde{u}}(y'|x', u')| \leq \|(y, x, u) - (y', x', u')\|_\infty,$$
$$\forall (y, x, u), (y', x', u') \in C_{\tilde{y},\tilde{x},\tilde{u}}.$$

Thus, $E_{\tilde{y},\tilde{x},\tilde{u}}$ satisfies the Lipschitz continuity Assumption 2.1, with Lipschitz constant $K = 1$, on the set $C_{\tilde{y},\tilde{x},\tilde{u}}$. The function $E_{\tilde{y},\tilde{x},\tilde{u}}$ is continuous at the boundary of $C_{\tilde{y},\tilde{x},\tilde{u}}$ and is zero outside $C_{\tilde{y},\tilde{x},\tilde{u}}$. Thus, $E_{\tilde{y},\tilde{x},\tilde{u}}$ is obtained by piecing together in a continuous manner a Lipschitz continuous function and a constant function. It follows that $E_{\tilde{y},\tilde{x},\tilde{u}}$ is Lipschitz continuous on the set $S \times S \times U$, with Lipschitz constant 1.

We define an instance $(P, g, \alpha, \varepsilon)$ by letting

$$g(x, u) = x_1, \qquad \forall (x, u) \in S \times U, \tag{3.4}$$

and

$$P(y|x, u) = 1, \qquad \forall (y, x, u) \in S \times S \times U. \tag{3.5}$$

It is easily seen that this instance satisfies Assumptions 2.1–2.3 with $K = 1$ and $\rho = 1/2$.

Bellman's equation reads

$$J(x) = x_1 + \alpha \int_S J(y)\, dy. \tag{3.6}$$

A simple calculation shows that the function J^* defined by

$$J^*(x) = x_1 + \frac{\alpha}{2(1 - \alpha)} \tag{3.7}$$

is a solution of (3.6) and according to the discussion of Section 2, it is the unique fixed point of T.

Let X be the set of points (y, x, u) sampled by the particular algorithm we are considering, when it is faced with the instance $(P, g, \alpha, \varepsilon)$. In particular, the cardinality of X is at most A. Using the definition of δ [cf. Eqs. (3.1) and (3.2)], the cardinality of X is at most $1/(4\delta^{2n+m})$.

We say that a cell $C_{\tilde{y},\tilde{x},\tilde{u}}$ is *sampled* if the intersection of X and $C_{\tilde{y},\tilde{x},\tilde{u}}$ is nonempty. Otherwise, we say that $C_{\tilde{y},\tilde{x},\tilde{u}}$ is *unsampled*. We say that some $(\tilde{x}, \tilde{u}) \in \tilde{S} \times \tilde{U}$ is *well-sampled* if there exist at least $1/(2\delta^n)$ elements $\tilde{y}$ of $\tilde{S}$ for which the cell $C_{\tilde{y},\tilde{x},\tilde{u}}$ is sampled. Otherwise, we say that $(\tilde{x}, \tilde{u})$ is *badly sampled*. Since the total number of samples is bounded by $1/(4\delta^{2n+m})$, there exist at most $1/(2\delta^{n+m})$ well-sampled elements $(\tilde{x}, \tilde{u}) \in \tilde{S} \times \tilde{U}$. Therefore, there are at least $1/(2\delta^{n+m})$ badly sampled $(\tilde{x}, \tilde{u})$. For each $\tilde{x} \in \tilde{S}$ there are at most $1/\delta^m$ possible choices of $\tilde{u}$ such that $(\tilde{x}, \tilde{u})$ is badly sampled. This shows that there exists a set $\tilde{S}_{\mathrm{BAD}} \subset \tilde{S}$ of cardinality $1/(2\delta^n)$ such that for each $\tilde{x} \in \tilde{S}_{\mathrm{BAD}}$ there exists some $\tilde{\mu}(\tilde{x}) \in \tilde{U}$ for which $(\tilde{x}, \tilde{\mu}(\tilde{x}))$ is badly sampled.

We will now construct a second instance. The cost function g is left unchanged [cf. Eq. (3.4)], but we modify the probability density on some of the unsampled cells. This is done as follows. Let us fix some $\tilde{x} \in \tilde{S}_{\mathrm{BAD}}$. By the definition of $\tilde{S}_{\mathrm{BAD}}$ and $\tilde{\mu}(\tilde{x})$, if we keep $\tilde{x}$ fixed and vary $\tilde{y}$, we find at least $1/(2\delta^n)$ unsampled cells of the form $C_{\tilde{y},\tilde{x},\tilde{\mu}(\tilde{x})}$. We sort these unsampled cells in order of increasing $\tilde{y}_1$ and we let c be the median value of $\tilde{y}_1$. We refer to those cells for which $\tilde{y}_1 \le c$ (respectively, $\tilde{y}_1 \ge c$) as *low* (respectively, *high*) cells. Let $\underline{c} = c - (1/16)$ and $\bar{c} = c + (1/16)$. We discard all unsampled cells $C_{\tilde{y},\tilde{x},\tilde{\mu}(\tilde{x})}$ for which $\underline{c} < \tilde{y}_1 < \bar{c}$. Thus, the number of discarded cells is bounded by $1/(8\delta^n)$. Since we started with at least $1/(4\delta^n)$ low unsampled cells, we are left with at least $1/(8\delta^n)$ such cells. By discarding some more low unsampled cells (if needed), we can assume that we are left with exactly $1/(8\delta^n)$ unsampled low cells. By a similar argument, we can also assume that we are left with exactly $1/(8\delta^n)$ unsampled high cells. Let $Q^{\mathrm{L}}(\tilde{x})$ [respectively, $Q^{\mathrm{H}}(\tilde{x})$] be the set of all $\tilde{y} \in \tilde{S}$ such that $C_{\tilde{y},\tilde{x},\tilde{\mu}(\tilde{x})}$ is a low (respectively, high) unsampled cell that has not been discarded. This procedure is carried out for each $\tilde{x} \in \tilde{S}_{\mathrm{BAD}}$.

We define

$$\hat{P}(y|x, u) = P(y|x, u) + E(y|x, u) = 1 + E(y|x, u), \tag{3.8}$$

where

$$E(y|x, u) = \sum_{\tilde{x}\in\tilde{S}_{\mathrm{BAD}},\, \tilde{y}\in Q^{\mathrm{L}}(\tilde{x})} E_{\tilde{y},\tilde{x},\tilde{\mu}(\tilde{x})}(y|x, u) - \sum_{\tilde{x}\in\tilde{S}_{\mathrm{BAD}},\, \tilde{y}\in Q^{\mathrm{H}}(\tilde{x})} E_{\tilde{y},\tilde{x},\tilde{\mu}(\tilde{x})}(y|x, u). \tag{3.9}$$

In words, we add a pyramid at each low unsampled cell and we subtract a pyramid at each high unsampled cell. This has the effect of shifting the transition probability distribution closer to the origin, with a consequent decrease in the cost incurred after a transition.

We verify that our perturbed instance $(\hat{P}, g, \alpha, \varepsilon)$ satisfies the required assumptions. Since each pyramid is Lipschitz continuous with Lipschitz constant 1, and since distinct pyramids are supported on distinct cells, it follows that Assumption 2.1(b) is satisfied with $K = 1$. Furthermore, for each (x, u), the number of added pyramids is equal to the number of substracted pyramids. For this reason, $\int_S E(y|x, u)\,dy = 0$ and $\hat{P}$ satisfies Assumption 2.2. Finally, the height of each pyramid is $\delta/2$. Since $\delta \leq 1$, we have $\hat{P}(y|x, u) \geq 1 - (\delta/2) \geq 1/2$. This shows that Assumption 2.2 and Assumption 2.3 (with $\rho = 1/2$) are satisfied.

Our next task is to estimate the optimal cost function $\hat{J}^*$ corresponding to the perturbed instance $(\hat{P}, g, \alpha, \varepsilon)$. Let

$$B = \left\{x \in S \mid \exists \tilde{x} \in \tilde{S}_{\mathrm{BAD}} \text{ such that } \|x - \tilde{x}\|_\infty \leq \frac{\delta}{4}\right\}.$$

For any $x \in B$, we let $\mu(x) = \tilde{\mu}(\tilde{x})$, where $\tilde{x}$ is the element of $\tilde{S}_{\mathrm{BAD}}$ for which $\|x - \tilde{x}\|_\infty \leq \delta/4$. For any $x \notin B$, we let $\mu(x) = 0$. We now consider the quantity

$$e(x) = \int_S g(y)E(y|x, \mu(x))\,dy, \tag{3.10}$$

which can be interpreted as the effect of the perturbation on the expected cost after the first transition, when the control is chosen according to the function μ.

LEMMA 3.1. *For each $x \in S$, we have $e(x) \leq 0$. Furthermore, there exists a positive absolute constant k such that $e(x) \leq -k\delta$, for all $x \in B$.*

Proof. Using Eqs. (3.9) and (3.10) and the definition of g, we have

$$\begin{aligned} e(x) = &\sum_{\tilde{x} \in \tilde{S}_{\mathrm{BAD}},\, \tilde{y} \in Q^{\mathrm{L}}(\tilde{x})} \int_S y_1 E_{\tilde{y}, \tilde{x}, \tilde{\mu}(\tilde{x})}(y|x, \mu(x))\,dy \\ &- \sum_{\tilde{x} \in \tilde{S}_{\mathrm{BAD}},\, \tilde{y} \in Q^{\mathrm{H}}(\tilde{x})} \int_S y_1 E_{\tilde{y}, \tilde{x}, \tilde{\mu}(\tilde{x})}(y|x, \mu(x))\,dy. \end{aligned} \tag{3.11}$$

For any $x \notin B$, we have $\mu(x) = 0$, which implies that $E(y|x, \mu(x)) = 0$ and $e(x) = 0$. Let us now fix some $x \in B$ and let $\tilde{x}$ be the corresponding element of $\tilde{S}_{\mathrm{BAD}}$. Then, Eq. (3.11) becomes

$$e(x) = \sum_{\tilde{y}\in Q^{\mathrm{L}}(\tilde{x})} \int_S y_1 E_{\tilde{y},\tilde{x},\tilde{\mu}(\tilde{x})}(y|x, \mu(x))\, dy$$

$$- \sum_{\tilde{y}\in Q^{\mathrm{H}}(\tilde{x})} \int_S y_1 E_{\tilde{y},\tilde{x},\tilde{\mu}(\tilde{x})}(y|x, \mu(x))\, dy. \tag{3.12}$$

Let us consider the summand corresponding to a particular $\tilde{y} \in Q^{\mathrm{L}}(\tilde{x})$. We need only carry out the integration on the set $Y(\tilde{y}) = \{y \in S \mid \|y - \tilde{y}\|_\infty \leq \delta/2\}$ (instead of the entire set S) because $E_{\tilde{y},\tilde{x},\tilde{\mu}(\tilde{x})}(y|x, u)$ vanishes when $y \notin Y(\tilde{y})$. For $y \in Y(\tilde{y})$, we have $y_1 \leq \tilde{y}_1 + \delta/2 \leq \underline{c} + \delta/2$. We now use the definition of the function $E_{\tilde{y},\tilde{x},\tilde{\mu}(\tilde{x})}(y|x, \mu(x))$ [cf. Eq. (3.3)], together with the property $\mu(x) = \tilde{\mu}(\tilde{x})$, to conclude that

$$\int_S y_1 E_{\tilde{y},\tilde{x},\tilde{\mu}(\tilde{x})}(y|x, \mu(x))\, dy \leq \left(\underline{c} + \frac{\delta}{2}\right) I(x),$$

where

$$I(x) = \int_{Y(\tilde{y})} \left(\frac{\delta}{2} - \max\{\|x - \tilde{x}\|_\infty, \|y - \tilde{y}\|_\infty\}\right) dy. \tag{3.13}$$

It is clear that the value of $I(x)$ is independent of the choice of $\tilde{y}$, which justifies our notation. By a symmetrical argument, each one of the summands corresponding to $\tilde{y} \in Q^{\mathrm{H}}(\tilde{x})$ is bounded below by $(\bar{c} - \delta/2)I(x)$. Since each one of the sets $Q^{\mathrm{H}}(\tilde{x})$, $Q^{\mathrm{L}}(\tilde{x})$ has cardinality $1/(8\delta^n)$, it follows from Eq. (3.11) that

$$e(x) \leq -(\bar{c} - \underline{c} - \delta)I(x)\frac{1}{8\delta^n} \leq -I(x)\frac{1}{128\delta^n}, \qquad \forall x \in B, \tag{3.14}$$

where the last inequality follows because $\bar{c} - \underline{c} = 1/8$ (by construction) and $\delta \leq 1/16$ (by definition). We now bound $I(x)$ for $x \in B$. We have $\|x - \tilde{x}\|_\infty \leq \delta/4$, and the integrand is always nonnegative and is at least $\delta/4$ for every y belonging to the set $\{y \in S \mid \|y - \tilde{y}\|_\infty \leq \delta/4\}$. Therefore, for $x \in B$, $I(x)$ is bounded below by $\delta/4$ times the volume of the set $\{y \in S \mid \|y - \tilde{y}\|_\infty \leq \delta/4\}$. This set is an n-dimensional cube, whose edges have length $\delta/2$. Thus, we obtain

$$I(x) \geq \frac{\delta}{4}\frac{\delta^n}{2^n}, \qquad \forall x \in B.$$

Combining with Eq. (3.14), we obtain

$$e(x) \leq -\frac{\delta}{4 \cdot 128 \cdot 2^n}, \qquad \forall x \in B,$$

which proves the desired result. Q.E.D.

LEMMA 3.2. *There exists a positive absolute constant h such that*

$$\int_B \hat{P}(y|x, u)\, dy \geq h, \qquad \forall x \in S, \forall u \in U.$$

Proof. The function $\hat{P}$ is bounded below by 1/2, as shown earlier. Thus, it suffices to show that the volume of B is bounded below by some absolute constant. Note that B consists of $1/(2\delta^n)$ cubes of volume $(\delta/2)^n$, and the result follows. Q.E.D.

Let $\hat{T}$ be the operator defined by Eq. (2.1) but with P replaced by $\hat{P}$. We have

$$\begin{aligned}
\hat{T}J^*(x) &= g(x) + \alpha \min_{u \in U} \int_S J^*(y)\hat{P}(y|x, u)\, dy \\
&\leq g(x) + \alpha \int_S J^*(y)\hat{P}(y|x, \mu(x))\, dy \\
&= x_1 + \alpha \int_S J^*(y)\, dy + \alpha \int_S J^*(y)E(y|x, \mu(x))\, dy \qquad (3.15) \\
&= J^*(x) + \alpha \int_S J^*(y)E(y|x, \mu(x))\, dy \\
&= J^*(x) + \alpha \int_S y_1 E(y|x, \mu(x))\, dy + \frac{\alpha^2}{2(1-\alpha)} \int_S E(y|x, \mu(x))\, dy \\
&= J^*(x) + \alpha e(x),
\end{aligned}$$

where we have used the fact that J^* satisfies Eqs. (3.6) and (3.7), the definition of $e(x)$ [cf. Eq. (3.10)], and the fact that $\int_S E(y|x, \mu(x))\, dy = 0$. It follows that

$$\hat{T}J^*(x) \leq J^*(x), \qquad \forall x \in S, \tag{3.16}$$

$$\hat{T}J^*(x) \leq J^*(x) - \alpha k \delta, \qquad \forall x \in B, \tag{3.17}$$

where k is the constant of Lemma 3.1. Let $\hat{T}^t$ be the composition of t copies of $\hat{T}$ and let $\bar{B} = \{x \in S \mid x \notin B\}$ be the complement of B. We have

$$
\begin{aligned}
\hat{T}^2 J^*(x) &= g(x) + \alpha \min_{u\in U} \int_S \hat{T}J^*(y)\hat{P}(y|x, u)\, dy \\
&\le g(x) + \alpha \int_S \hat{T}J^*(y)\hat{P}(y|x, \mu(x))\, dy \\
&= g(x) + \alpha \int_B \hat{T}J^*(y)\hat{P}(y|x, \mu(x))\, dy \\
&\quad + \alpha \int_{\bar{B}} \hat{T}J^*(y)\hat{P}(y|x, \mu(x))\, dy \qquad (3.18)\\
&\le g(x) + \alpha \int_B (J^*(y) - \alpha k\delta)\hat{P}(y|x, \mu(x))\, dy \\
&\quad + \alpha \int_{\bar{B}} J^*(y)\hat{P}(y|x, \mu(x))\, dy \\
&= J^*(x) + \alpha e(x) - \alpha^2 k\delta \int_B \hat{P}(y|x, \mu(x))\, dy \\
&\le J^*(x) - \alpha^2 k\delta h, \qquad \forall x \in S.
\end{aligned}
$$

[We have used here the equality between the second and the last line of Eq. (3.15), as well as Lemma 3.2.] It is well known (Bertsekas, 1987) (and easy to verify) that for any real constant d, we have $\hat{T}(J + d) = \alpha d + \hat{T}J$. (Here the notation $J + d$ should be interpreted as the function which is equal to the sum of J with a function on S that is identically equal to d.) Using this property and Eq. (3.18), we obtain

$$\hat{T}^3 J^*(x) \le \hat{T}^2 J^*(x) - \alpha^3 k\delta h \le J^*(x) - \alpha^2 k\delta h - \alpha^3 k\delta h, \qquad \forall x \in S.$$

We continue inductively, to obtain

$$\hat{T}^t J^*(x) \le J^*(x) - (1 + \alpha + \alpha^2 + \cdots + \alpha^{t-2})\alpha^2 k\delta h, \qquad t = 2, 3, \ldots, \forall x \in S. \qquad (3.19)$$

Taking the limit as $t \to \infty$, $\hat{T}^t J^*$ converges to the optimal cost function $\hat{J}^*$ of the perturbed instance, and Eq. (3.19) implies that

$$\hat{J}^*(x) \le J^*(x) - \frac{\alpha^2}{1 - \alpha}\delta k h, \qquad \forall x \in S. \qquad (3.20)$$

Note that the perturbed instance coincides with the original one at all points sampled by the algorithm. For this reason, the algorithm will perform the same arithmetic operations and will return the same answer for both instances. That answer must be an ε-approximation of both J^* and $\hat{J}^*$. It follows that $\|J^* - \hat{J}^*\|_\infty \leq 2\varepsilon$. Therefore,

$$\frac{\alpha^2}{1-\alpha}\,\delta k h \leq 2\varepsilon.$$

Since $\alpha \geq 1/2$, we obtain

$$\delta \leq d(1-\alpha)\varepsilon, \tag{3.21}$$

where d is some absolute constant.

For $\varepsilon \leq 1/(32d)$, we obtain $\delta \leq 1/32$ or $1/(2\delta) \geq 16$. Thus, using Eq. (3.1), we have $1/\delta_0 \geq (1/\delta) - 16 \geq (1/\delta) - (1/2\delta) = 1/(2\delta)$, and Eq. (3.2) yields

$$A = \frac{1}{4\delta_0^{2n+m}} \geq \frac{1}{4(2\delta)^{2n+m}} \geq \frac{1}{4(2d(1-\alpha)\varepsilon)^{2n+m}} = \Omega\left(\frac{1}{((1-\alpha)\varepsilon)^{2n+m}}\right).$$

When the theorem is proved for general values of K and ρ, it is sufficient to multiply the pyramidal functions of Eq. (3.3) by a factor of $\min\{K, 1-\rho\}$. It is then easily seen that the perturbed problem satisfies Assumptions 2.1 and 2.3 for the given values of K and ρ and the proof goes through verbatim, except that certain absolute constants are modified. Q.E.D.

In our next result, the mixing condition (Assumption 2.3) is removed. It will be seen that this allows us to obtain a larger lower bound.

THEOREM 3.2 (lower bound under Assumptions 2.1 and 2.2). *For every m, n, there exists some K such that*

$$\mathbf{C}_{\text{prob}}(\varepsilon, \alpha) = \Omega\left(\frac{1}{((1-\alpha)^2\varepsilon)^{2n+m}}\right).$$

Proof. The structure of the proof is similar to that of the preceding proof. We fix n, m, and some K that will depend on n in a way to be determined later. An absolute constant is again a constant that depends only on m and n.

We fix some $\varepsilon > 0$ and some $\alpha \in (1/2, 1)$. We consider an algorithm that is correct for the problem $\mathcal{P}_{\text{prob}}$ (for the given values of m, n, K) and suppose that the number of queries is at most A for every instance with those particular values of α and ε.

We choose a positive scalar δ so that $1/\delta$ is an integer multiple of 9 and such that

$$\frac{1}{\delta_0} \leq \frac{1}{\delta} \leq \frac{1}{\delta_0} + 9, \tag{3.22}$$

where δ_0 satisfies

$$A = \frac{1}{5}\left(\frac{2}{9}\right)^{2n} \frac{1}{\delta_0^{2n+m}}. \tag{3.23}$$

We partition $S \times S \times U$ into cubic cells of volume $1/\delta^{2n+m}$ exactly as in the proof of Theorem 3.1 and we use the same notations, $\tilde{S}$, $\tilde{U}$, $C_{\tilde{y},\tilde{x},\tilde{u}}$, and $E_{\tilde{y},\tilde{x},\tilde{u}}$.

We define the first instance to be considered. Let F_1, F_2, G: $[0, 1] \mapsto \Re$ be the functions shown in Fig. 1. We define a function H: $[0, 1] \times [0, 1] \mapsto \Re$ by letting

$$H(y|x) = F_1(y)G(x) + F_2(y)(1 - G(x)), \qquad \forall x, y \in [0, 1].$$

We finally let

$$P(y|x, u) = \prod_{i=1}^{n} H(y_i|x_i), \qquad \forall (y, x, u) \in S \times S \times U, \tag{3.24}$$

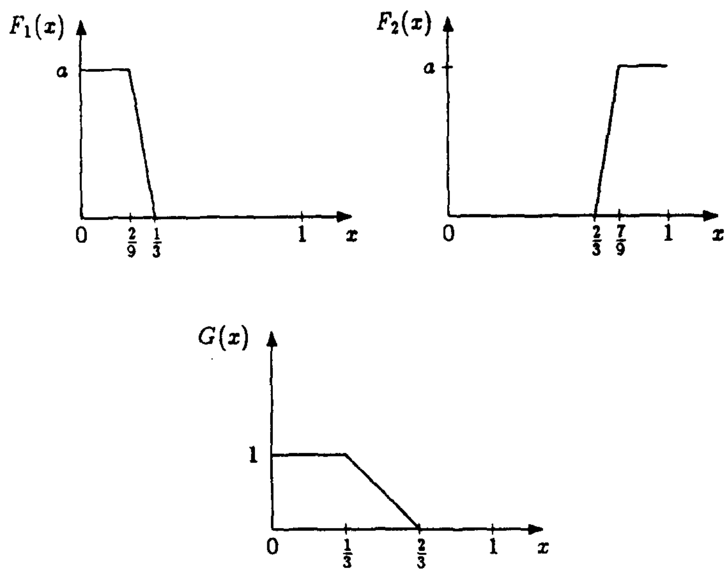

FIG. 1. The functions F_1, F_2, and G. The maximum value a of F_1 and F_2 is chosen so that $\int_0^1 F_1(x)\,dx = \int_0^1 F_2(x)\,dx = 1$. In particular, $3 \leq a \leq 9/2$.

where x_i and y_i are the ith components of x and y, respectively. As for the cost function g, we only assume that $g(x, u) = 1$ for all $x \in [0, 1/3]^n$ and $u \in U$, and that $g(x, u) = 0$ for all $x \in [2/3, 1]^n$ and $u \in U$.

We verify that Assumptions 2.1 and 2.2 are satisfied. The function P is certainly nonnegative. Furthermore, F_1 and F_2 integrate to 1. Consequently, $\int_{[0,1]} H(y|x)\, dy = 1$, for all $x \in [0, 1]$. Thus, for any x, u, $P(\cdot|x, u)$ is a product of probability measures [cf. Eq. (3.24)] and is itself a probability measure. Note that F_1, F_2, and G are Lipschitz continuous. It follows that P is also Lipschitz continuous with Lipschitz constant K, provided that the absolute constant K is taken large enough. Concerning the function g, we have not specified it in detail, but it is easily seen that there exist Lipschitz continuous functions satisfying the requirements we have imposed on g.

Note that the Markov chain corresponding to P has the property that if the current state is in the set $[0, 1/3]^n$ then the state stays forever in that set. The same property holds for the set $[2/3, 1]^n$.

We now estimate $J^*(x)$ when $x \in [0, 1/3]^n$. While we could argue directly in terms of the Bellman equation, the argument is much more transparent if we use the interpretation of $J^*(x)$ as the optimal cost expressed as a function of the initial state. Starting with some initial state in $[0, 1/3]^n$, the state never exits that set. Furthermore, $g(x) = 1$ for every $x \in [0, 1/3]^n$. This implies that

$$J^*(x) = \sum_{t=0}^{\infty} \alpha^t = \frac{1}{1-\alpha}, \qquad \forall x \in [0, 1/3]^n. \tag{3.25}$$

LEMMA 3.3. *There exists a set* $\tilde{S}_{\mathrm{BAD}} \subset [0, 2/9]^n \cap \tilde{S}$ *of cardinality* $(2/9)^n/(2\delta^n)$ *with the following property: for every* $\tilde{x} \in \tilde{S}_{\mathrm{BAD}}$ *there exist some* $\tilde{\mu}(\tilde{x}) \in \tilde{U}$ *and two sets* $Q^{\mathrm{L}}(\tilde{x}) \subset [0, 2/9]^n \cap \tilde{S}$, $Q^{\mathrm{H}}(\tilde{x}) \subset [7/9, 1]^n \cap \tilde{S}$, *each of cardinality* $(2/9)^n/(2\delta^n)$, *such that the cell* $C_{\tilde{y},\tilde{x},\tilde{\mu}(\tilde{x})}$ *is unsampled for every* $\tilde{y} \in Q^{\mathrm{L}}(\tilde{x}) \cup Q^{\mathrm{H}}(\tilde{x})$.

Proof. Let $\tilde{S}_{\mathrm{GOOD}}$ be the set of all $\tilde{x} \in [0, 2/9]^n \cap \tilde{S}$ that do not have the desired property. Since the cardinality of $[0, 2/9]^n \cap \tilde{S}$ is $(2/9)^n/\delta^n$ it is sufficient to show that $\tilde{S}_{\mathrm{GOOD}}$ has cardinality less than or equal to $(2/9)^n/(2\delta^n)$. We suppose the contrary, and we will obtain a contradiction.

Fix some $\tilde{x} \in \tilde{S}_{\mathrm{GOOD}}$. Then, for every $\tilde{u} \in \tilde{U}$ we can find at least $(2/9)^n/(2\delta^n)$ values of $\tilde{y} \in \tilde{S}$ such that the cell $C_{\tilde{y},\tilde{x},\tilde{u}}$ is sampled. This shows that the total number of sampled cells is at least $(2/9)^{2n}/(4\delta^{2n+m})$. Using Eqs. (3.22) and (3.23), this implies that the number of sampled cells is more than A, a contradiction. Q.E.D.

We now construct a perturbed instance. The cost function g is left unchanged. We define

$$\hat{P}(y|x, u) = P(y|x, u) + E(y|x, u),$$

where

$$E(y|x, u) = \sum_{\tilde{x}\in\tilde{S}_{\mathrm{BAD}},\, \tilde{y}\in Q^{\mathrm{H}}(\tilde{x})} E_{\tilde{y},\tilde{x},\tilde{\mu}(\tilde{x})}(y|x, u) - \sum_{\tilde{x}\in\tilde{S}_{\mathrm{BAD}},\, \tilde{y}\in Q^{\mathrm{L}}(\tilde{x})} E_{\tilde{y},\tilde{x},\tilde{\mu}(\tilde{x})}(y|x, u). \tag{3.26}$$

In effect, we are giving positive probability to certain transitions from the set $[0, 2/9]^n$ to the set $[7/9, 1]^n$. On the other hand, the property that the state can never exit from the set $[7/9, 1]^n$ is retained. The Lipschitz continuity of E and P implies that $\hat{P}$ is Lipschitz continuous. Also, $\hat{P}(\cdot|x, u)$ is nonnegative and integrates to 1, for reasons similar to those in the proof of Theorem 3.1. Thus, Assumptions 2.1 and 2.2 are satisfied.

Let

$$B = \left\{x \in [0, 1/3]^n \mid \exists \tilde{x} \in \tilde{S}_{\mathrm{BAD}} \text{ such that } \|x - \tilde{x}\|_\infty \le \frac{\delta}{4}\right\}.$$

LEMMA 3.4. *For every $x \in [0, 1/3]^n$, we have*

$$\int_B \hat{P}(y|x, u)\, dy \ge h,$$

where h is a positive absolute constant.

Proof. Fix some $x \in [0, 1/3]^n$. Note that $P(y|x, u) = \prod_{i=1}^n F_2(y_i) \ge 3^n \ge 3$ for all $y \in [0, 2/9]^n$. Since $|E(y|x, u)| \le \delta/2 \le 1$, we conlude that $\hat{P}(y|x, u) \ge 2$, for all $y \in [0, 2/9]^n$. The set B consists of $(2/9\delta)^n/2$ cubes of volume $(\delta/2)^n$. Thus, the volume of B is bounded below by some absolute positive constant, and the result follows. Q.E.D.

Let us now define $\mu(x) = \tilde{\mu}(\tilde{x})$ for all $\tilde{x} \in B$, where $\tilde{x}$ is chosen so that $\|x - \tilde{x}\|_\infty \le \delta/4$, and we let $\mu(x) = 0$ for $x \notin B$.

LEMMA 3.5. *For every $x \in B$, we have*

$$\int_{[2/3,1]^n} \hat{P}(y|x, \mu(x))\, dy \ge k\delta,$$

where k is an absolute positive constant.

Proof. For every $\tilde{y} \in \tilde{S}$, let $Y(\tilde{y}) = \{y \in S \mid \|y - \tilde{y}\|_\infty \leq \delta/4\}$. Fix some $x \in B$ and let $\tilde{x}$ be an element of $\tilde{S}_{\text{BAD}}$ such that $\|x - \tilde{x}\|_\infty \leq \delta/4$. We have

$$\int_{[2/3,1]^n} \hat{P}(y|x, \mu(x))\, dy$$

$$\geq \sum_{\tilde{y}\in Q^{\mathrm{H}}(\tilde{x})} \int_{Y(\tilde{y})} E_{\tilde{y},\tilde{x},\tilde{\mu}(\tilde{x})}(y|x, \mu(x))\, dy$$

$$\geq \sum_{\tilde{y}\in Q^{\mathrm{H}}(\tilde{x})} \int_{Y(\tilde{y})} \left(\frac{\delta}{2} - \max\{\|x - \tilde{x}\|_\infty, \|y - \tilde{y}\|_\infty\}\right) dy$$

$$\geq \sum_{\tilde{y}\in Q^{\mathrm{H}}(\tilde{x})} \frac{\delta}{4} \int_{Y(\tilde{y})} dy.$$

The set $Y(\tilde{y})$ is a cube of volume $(\delta/2)^n$, the cardinality of the set $Q^{\mathrm{H}}(\tilde{x})$ is $(2/9\delta)^n/2$, and the result follows. Q.E.D.

We now estimate the cost $J_\pi(x)$ which is incurred if policy $\pi = (\mu, \mu, \ldots)$ is used, for the case where $x \in [0, 1/3]^n$. The corresponding Markov process x_t^π evolves as follows. Whenever $x_t^\pi \in [0, 1/3]^n$, there is at least probability h that the next state belongs to the set B and there is a further probability of at least $k\delta$ that the state after one more transition is in the set $[2/3, 1]^n$. Once the latter set is entered, the state stays forever in that set. We therefore have

$$\Pr(x_t^\pi \in [0, 1/3]^n) \leq (1 - kh\delta)^{t-1}, \qquad \forall t \geq 1.$$

Since the cost is 1 on the set $[0, 1/3]^n$ and 0 on the set $[2/3, 1]^n$, we have

$$\begin{aligned} J_\pi(x) &= \sum_{t=0}^{\infty} \alpha^t \Pr(x_t^\pi \in [0, 1/3]^n) \\ &\leq 1 + \sum_{t=1}^{\infty} \alpha^t (1 - kh\delta)^{t-1} \qquad (3.27) \\ &= 1 + \frac{\alpha}{1 - \alpha(1 - kh\delta)} \\ &= \frac{1 + \alpha kh\delta}{1 - \alpha(1 - kh\delta)}, \qquad \forall x \in [0, 1/3]^n. \end{aligned}$$

The optimal cost function $\hat{J}^*$ of the perturbed instance satisfies $\hat{J}^* \leq \hat{J}_\pi$ and, using Eq. (3.25), we obtain

$$J^*(x) - \hat{J}^*(x) \geq \frac{1}{1-\alpha} - \frac{1 + \alpha kh\delta}{1 - \alpha(1 - kh\delta)}$$
$$= \frac{\alpha^2 kh\delta}{(1-\alpha)(1-\alpha(1-kh\delta))}, \qquad \forall x \in [0, 1/3]^n. \quad (3.28)$$

Note that the class $\mathcal{P}_{\text{prob}}$ contains the class $\mathcal{P}_{\text{mix}}$. For this reason, the particular algorithm being considered here is also a correct algorithm for the problem $\mathcal{P}_{\text{mix}}$. In particular, all of the intermediate results in the proof of Theorem 3.1 apply to the algorithm we are considering. We can therefore use Eq. (3.21) and conclude that $\delta \leq d(1 - \alpha)\varepsilon$, where d is an absolute constant. (Actually, the definition of δ is somewhat different in the two proofs, but this affects only the absolute constant d.) This implies that for $\varepsilon \leq 1/(khd)$, we have $\delta \leq (1-\alpha)/(kh)$ or $1 - kh\delta \geq \alpha$. Using this inequality in Eq. (3.28), together with the property $\alpha \geq 1/2$, we obtain

$$J^*(x) - \hat{J}^*(x) \geq \frac{\alpha^2 kh\delta}{(1-\alpha)(1-\alpha^2)} \geq \frac{(kh\delta)/4}{(1-\alpha)^2 2} = \frac{1}{8}\frac{kh\delta}{(1-\alpha)^2}.$$

This inequality is similar to inequality (3.20) in the proof of Theorem 3.1, except that $1 - \alpha$ has been replaced by $(1-\alpha)^2$. The rest of the argument is the same, except for certain constant factors and the fact that $1 - \alpha$ is replaced throughout by $(1-\alpha)^2$. Q.E.D.

THEOREM 3.3 (lower bound under Assumptions 2.1, 2.3, 2.4). *For every m, n, there is a choice of K and ρ such that*

$$\mathbf{C}_{\text{sub}}(\varepsilon, \alpha) = \Omega\left(\frac{1}{((1-\alpha)^2\varepsilon)^{2n+m}}\right).$$

Proof. The proof is almost identical to the proof of Theorem 3.2, and for this reason, we argue informally. For convenience, let the state space S be the set $[0, 1/3]^n$, instead of $[0, 1]^n$, and let $P(y|x, u)$ be defined on that set as in the proof of Theorem 3.2. Then, P is a probabilty measure on the set $[0, 1/3]^n$ and the corresponding function J^* is identically equal to $1/(1-\alpha)$. Note that P satisfies Assumption 2.3. Let $\hat{P}$ be as in the proof of Theorem 3.2, except that it is defined only for $x, y \in [0, 1/3]^n$. For this reason, $\hat{P}$ is now a subprobability measure. The function $\hat{J}^*$ for the current problem is equal to the optimal expected discounted cost until the termination of the stochastic process. However, the process considered here terminates exactly when the process considered in the proof of Theorem 3.2 makes a transition from $[0, 1/3]^n$ to the zero-cost set $[2/3, 1]^n$. For this reason, the function $\hat{J}^*$ is the same as the function $\hat{J}^*$ in the proof of Theorem 3.2, and the result follows with the same reasoning. Q.E.D.

Remarks

1. Suppose that we replace the correctness requirement $\|J - J^*\|_\infty \le \varepsilon$ by the requirement $\|J - J^*\|_p \le \varepsilon$, where $1 \le p < \infty$ and $\|\cdot\|_p$ is the usual L_p-norm. Then, Theorems 3.1–3.3 remain true, with exactly the same proofs. The reason is that in all of our proofs we have constructed our perturbed instances so that $J^*(x) - \hat{J}^*(x)$ is "large" on a set whose measure is bounded below by an absolute constant [cf. Eq. (3.20) or Eq. (3.28)]. But this implies that $J^* - \hat{J}^*$ is also large when measured by the L_p-norm and the proofs remain valid, except that certain constants must be changed.

2. The lower bounds of Theorem 3.3 can also be proved for all values of the constants K and ρ. The proof is similar except that we should let $P(y|x, u) = 3^n$ for all $(y, x, u) \in S \times S \times U$, so that P satisfies the Lipschitz continuity assumption for any value of K. Furthermore, the perturbing pyramids should be multiplied by a factor that ensures that their Lipschitz constant is less than K and that Assumption 2.3 is not violated.

3. We are not able to establish the lower bound of Theorem 3.2 for an arbitrary choice of K. There is a simple reason for that: if K is taken very small, then Assumption 2.3 is automatically satisfied and the best provable lower bound is the one in Theorem 3.1.

4. Given some $\varepsilon > 0$, we say that a function μ: $S \mapsto U$ is ε-optimal if $J^*(x) \le J_\pi(x) \le J^*(x) + \varepsilon$, for all $x \in S$, where $\pi = (\mu\ \mu, \ldots)$. If an ε-optimal function μ is available, then the function J^* is automatically determined within an error of ε, the error being measured according to a norm $\|\cdot\|_p$. Thus, the number of oracle queries needed for computing an ε-optimal function μ is at least as large as the number of queries needed to determine J^* within ε. It follows that the lower bounds of Theorems 3.1–3.3 also apply (under their respective assumptions) to the computation of ε-optimal functions μ.

4. Discussion

The lower bounds of Section 3 agree with the upper bounds of Section 2. Thus, we have completely characterized the number of queries needed for approximating J^*. This leaves the further question of evaluating the total complexity of approximating J^*, when arithmetic computations are taken into account. This issue is addressed by Chow and Tsitsiklis (1989). In particular, they introduce a multigrid version of the iterative algorithm $J := TJ$ and show that the total number of arithmetic operations and comparisons is

$$O\left(\frac{1}{((1-\alpha)\varepsilon)^{2n+m}}\right), \tag{4.1}$$

for the problem $\mathcal{P}_{\text{mix}}$, and

$$O\left(\frac{1}{((1-\alpha)^2\varepsilon)^{2n+m}} \cdot \frac{1}{|\log \alpha|}\right) = O\left(\frac{1}{((1-\alpha)^2\varepsilon)^{2n+m}} \cdot \frac{1}{1-\alpha}\right), \tag{4.2}$$

for the problems $\mathcal{P}_{\text{prob}}$ and $\mathcal{P}_{\text{sub}}$. Thus, for problem $\mathcal{P}_{\text{mix}}$, we have an optimal algorithm. For the problems $\mathcal{P}_{\text{prob}}$ and $\mathcal{P}_{\text{sub}}$, we are within a factor of $O(1/(1-\alpha))$ from the optimum. One might wish to close this gap but the prospects are not particularly bright because (a) there are no effective methods for proving lower bounds tighter than those provided by lower bounding the number of queries, and (b) it can be shown (Chow, 1989) that no algorithm in a certain family of multigrid methods can have complexity better than the one provided by Eq. (4.2).

We expect that our results can be extended to the case where bounds are imposed on second derivatives (more generally, derivatives of order r) of the functions P and g. Of course, the bounds should change, with the exponent $2n + m$ being replaced by a lower exponent, depending on r.

As mentioned in the Introduction, the case where the functions g and P do not depend on u (equivalently, the case where U is a singleton) makes the equation $J = TJ$ a linear Fredholm equation of the second kind. Our proofs and our results remain true, provided that the exponent m in our bounds is replaced by 0. In particular, if we let $n = 1$, our results agree with the results of Werschulz (1985).[2] Our results are different from those of Werschulz in a number of respects:

(a) We are not limited to the one-dimensional case.

(b) We quantify the dependence of the complexity on the parameter α, which is a measure of the ill-conditioning of the problem.

(c) On the other hand, unlike Werschulz (1985), we do not study the dependence of the complexity on the smoothness properties (e.g., bounds on higher derivatives) of the functions P and g.

References

Bertsekas, D. P. (1987), "Dynamic Programming: Deterministic and Stochastic Models," Prentice-Hall, Englewood Cliffs, NJ.

---

[2]Werschulz (1985) replaces the Lipschitz continuity assumption by a continuous differentiability assumption. Since the latter is a more restrictive assumption, our lower bounds do not apply to that case. However, the arguments in our proofs can be modified, by "smoothing" the corners and the edges of our "pyramids" and therefore the same lower bounds hold for the continuously differentiable case.

Bertsekas, D. P. (1975), Convergence of discretization procedures in dynamic programming, *IEEE Trans. Automat. Control* **AC-20,** 415–419.

Bertsekas, D. P., and Shreve, S. E. (1978), "Stochastic Optimal Control: The Discrete Time Case," Academic Press, New York.

Chow, C.-S. (1989), Ph.D. thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, in preparation.

Chow, C.-S., and Tsitsiklis, J. N. (1989), "An Optimal Multigrid Algorithm for Discrete-Time Stochastic Control," Technical Report LIDS-P-1864, Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA, April. (A preliminary version of this paper was presented at the 27th IEEE Conference on Decision and Control.)

Hartley, R., Thomas, L. C., and White, D. J. (Eds.) (1980), "Recent Developments in Markov Decision Processes," Academic Press, New York.

Nemirovsky, A. S., and Yudin, D. B. (1983), "Problem Complexity and Method Efficiency in Optimization," Wiley, New York.

Papadimitriou, C. H., and Tsitsiklis, J. N. (1987), The complexity of Markov decision processes, *Math. Oper. Res.* **12,** No. 3, 441–450.

Traub, J. F., Wasilkowski, G. W., and Woźniakowski, H. (1988), "Information-Based Complexity," Academic Press, San Diego, CA.

Werschulz, A. G. (1985), What is the complexity of the Fredholm problem of the second kind? *J. Integral Equations* **9,** 213–241.

Whitt, W. (1978), Approximations of dynamic programs, *Math. Oper. Res.* **3,** 231–243.